An Overview of the ISLSCP* Initiative I Global Data Sets

P.J. Sellers[1], B.W. Meeson[2]
J. Closs[3], J. Collatz[1], F. Corprew[3]
F.G. Hall[1], Y. Kerr[5], R. Koster[6], S. Los[7], K. Mitchell[8],
J. McManus[3], D. Myers[3], K.-J. Sun[3], P. Try[9]

1. NASA/GSFC, Code 923, Greenbelt, MD 20771
2. NASA/GSFC-DAAC, Code 902.2, Greenbelt, MD 20771
3. NASA/GSFC, HSTX, Code 902.2 Greenbelt, MD 20771
4. Colorado State University, Fort Collins, CO 80523
5. LERTS - BPI, Toulouse Cedex, 31055, France
6. NASA/GSFC, Code 974, Greenbelt, MD 20771
7. NASA/GSFC, SSAI, Code 923, Greenbelt, MD 20771
8. NOAA/NMC, Camp Springs, MD 20746
9. International GEWEX Project Office, Washington, DC20024

* International Satellite Land Surface Climatology Project

Abstract

In June of 1992, an interdisciplinary Earth Science workshop was convened in Columbia, Maryland, to assess recent progress in land-atmosphere research, specifically in the areas of models, satellite data algorithms, and field experiments. At the workshop, representatives of the land-atmosphere modeling community stated that they had a need for global data sets to prescribe boundary conditions, initialize state variables, and provide near-surface meteorological and radiative forcings for their models. The data sets collated on these CDs represent a first attempt to meet this need.

The data sets on the CDs are grouped under the following headings: Vegetation; Hydrology and Soils; Snow, Ice, and Oceans; Radiation and Clouds; and Near-Surface Meteorology.

All data sets cover the period 1987–1988, and all but a few are spatially continuous over tEarth's land surface. All have been mapped to a common 1° x 1° equal-angle grid. The temporal frequency for most of the data sets is monthly. A few of the near-surface meteorological parameters are available both as 6-hourly values and as monthly means.

CONTENTS

I. Background: The Motivation for Assembling the Initiative I Data Sets

A workshop sponsored by the International Satellite Land Surface Climatology Project (ISLSCP), a component of the Global Energy and Water Cycle Experiment (GEWEX), was held in Columbia, Maryland, June 23 to 26, 1992, with over 240 scientists and science managers attending. The goal of the workshop was to assess recent progress in the areas of modeling, satellite data algorithm development, and field experiments. An account of the workshop and many of the scientific presentations made there are written up in a special issue of the journal Remote Sensing of the Environment, 51, (1), 1995, so only a brief summary of the workshop's discussions and recommendations is given here.

1.1 Models

The first part of the workshop was spent in reviewing the goals and requirements set by different kinds of land-atmosphere models. For convenience, the models were categorized by time scale into three broad groups, see Sellers et al. (1995) and Figure 1.

Water-Energy-Carbon: These models are used to calculate the exchanges of water, energy, and carbon (photosynthesis and respiration) between the land surface and the atmosphere on relatively short time scales, on the order of seconds to seasons, see Dickinson (1995) and Bonan (1995). The models are used on small spatial scales for hydrological and agricultural studies; on the global scale, they are used to define the lower boundary fluxes for atmospheric general circulation models (GCMs) in which they are usually referred to as land surface parameterizations (LSPs). The more realistic biophysically based LSPs implemented within GCMs over the last decade have been shown to produce better simulations of energy and water fluxes over the continents and thus should give rise to improved numerical weather prediction and climate simulations, see Betts et al. (1994), Noilhan et al. (1991) and Sato et al. (1989). All of these models have suffered from two general weaknesses. First, it was not clear that the descriptions of important flux-controlling processes, e.g., heat and moisture transfer within the vegetation-soil-atmosphere system could be credibly transferred from models and observations tested or conducted at very small scales to the scales of LSP-GCMs. Second, no generally acceptable methods were available to define the global state of vegetation and soil moisture for initialization or validation of the LSP- GCMs.

Carbon and Biogeochemistry: The models appropriate to studies of carbon and biogeochemistry (BGC) span intermediate time scales, on the order of days to several years, see Schimel (1995) and Field et al. (1995). The important processes covered by these models include primary production, carbon allocation, decomposition, nutrient cycling and relations to the physical climate system (upstream), and ecosystem structure and function (downstream). It has been suggested that perturbations to the terrestrial carbon cycle, specifically imbalances between photosynthesis and respiration which would lead to carbon sink and source anomalies, may play important roles in the rate and timing of atmospheric carbon dioxide increases over the next few decades, see Tans et al. (1990). This class of models suffers from many of the same kind of handicaps as the Water Energy-Carbon models discussed above; in particular, global forcing (atmospheric conditions) and surface state (photosynthetic capacity, carbon storage in the soil, etc.) data are not freely available.

Ecosystem Structure and Function: These models have a large overlap with the carbon and BGC models but span a wider range of time scales; most of these models have time steps on the order of months to years and are run to describe ecological processes over periods of years to millennia, see Bonan (1995). In large part, the models are forced by climate data, but data on soil physical and chemical properties, topography, etc., are also used as boundary conditions. Obviously, all data pertaining to land cover type phenology, biomass, etc., are useful for initializing and validating these models.

The three classes of models described above originate from different scientific motivations and to a large extent from different science communities. However, they will all be essential for the study of Global Change and they all suffer from similar deficiencies, namely:

a. Scaling: All of the models suffer from the so-called scale gap to varying degrees. The results from small-scale process studies are usually combined with very simple aggregation assumptions to describe regional-scale processes and surface-atmosphere exchanges.

b. Data Needs: Very few reliable, consistent, large-scale data sets exist in accessible form for the purposes of initialization and validation of these models on regional or global scales.

In the first ISLSCP meetings of 1983 and 1984, it was hoped that a combination of large-scale field experiments and a stream of satellite data products would be used to deal with these two issues. Research work conducted within and parallel to the field experiments was to lead to improved algorithms which would then be used to generate better regional and global data sets. The 1992 ISLSCP workshop reviewed the state of the algorithms and the contribution of the field experiments to these goals.

1.2 Algorithms

Satellite data algorithms that deal with land surface studies have been developed piecemeal under the aegis of the responsible government agencies. In1987, an ISLSCP workshop reviewed the status of the algorithms, see Sellers et al. (1990) for a summary, and concluded that:

a. Algorithms were available to calculate many of the important surface and atmospheric state variables required by modelers.

b. Few of the algorithms had been thoroughly evaluated with regard to accuracy and precision.

c. There was a lot of room for improvement in the algorithms in terms of calibration, geometric correction, and cloud screening procedures.

d. Few of the algorithms had been tested sufficiently or were innately robust enough for routine operational use.

The 1992 workshop noted that there had been some progress in algorithm development over the period 1987–1992, particularly in the area of the Earth radiation budget work.

1.3 Field Experiments

The ISLSCP field experiments and parallel activities performed by the World Climate Research Program (WCRP), the International Geosphere-Biosphere Program (IGBP), and other organizations were designed to address the issues described in Section 1.1 above.

The results from several experiments, including FIFE, HAPEX-Mobilhy and others were presented and discussed at the meeting. In broad summary, the principal findings were as follows:

Scaling Issues: The problems of scaling soil-vegetation-atmosphere models from local scales up to several kilometers do not appear to be as severe as originally feared. This point was reemphasized at a joint ISLSCP-BAHC (BAHC; Biological Aspects of the Hydrological Cycle—an element of IGBP) workshop held in Tucson, Arizona, in March 1994, which focused specifically on scaling. A number of studies presented at the workshop indicated that the radiative transfer and mass and heat transport models used to describe processes on the scale of individual plants or small plots could be used to calculate large-scale (10–50 km) surface-atmosphere fluxes to acceptable accuracies using relatively simple spatial-aggregation techniques. In some of these studies, explicit checks were made on the accuracy of these methods using a variety of surface and airborne instruments to cover the scale range from a few centimeters out to several kilometers.

Use of Satellite Data: The field experiments sponsored by ISLSCP and other organizations involved the collection of integrated data sets, which allowed end-to-end evaluation of procedures for calculating surface state parameters from exoatmospheric radiances. It was found that several components of the surface radiation budget (insolation, downward photosynthetically active radiation (PAR), reflected shortwave) could be estimated from sensors on geostationary platforms to good accuracy and that useful estimates of downward longwave and net radiation could also be calculated. Satellite data were used to calculate surface biophysical parameters, including the fraction of photosynthetically active radiation absorbed by the green portion of the vegetation canopy (FPAR), unstressed stomatal conductance and photosynthetic capacity. These parameters have been used in simulation models to calculate the surface-atmosphere fluxes of carbon and water. Significantly, the remote sensing methodologies, the parameters, and the models themselves have been shown to be largely scale-invariant. This indicates that the local-scale models tested on the field experiment scale could be combined with large-scale satellite data sets to produce continental-scale fields of energy and mass (water and carbon dioxide) fluxes.

The field experiments succeeded in dealing with the two major issues that framed their design. The next task was to take the lessons learned from the experiments and apply them to improve models and to generate better data sets. It can be argued that the modeling community directly benefited from the work: several offline models and at least three operational GCMs currently utilize formulations that are based on field experiment results. However, it is also clear that the experimental results were only occasionally used to help generate improved large-scale data sets from satellite data.

1.4 The Need for Global Data Sets

It was concluded at the 1992 workshop that the communities working on model development and on past and planned field experiments had their activities in hand. However, it was made clear that the availability and accessibility of global data sets for land-atmosphere models were unsatisfactory.

Each modeling group reaffirmed the need for global data sets for initialization and boundary conditions, forcing, and validation, see Figure 2. The stated intention was to thoroughly test the surface models independent of atmospheric models, which cannot be relied on to provide realistic forcings, so as to highlight the components of the land models that need attention. Figure 2 shows the roles of the different data required to do this task. These are summarized below:

Surface Boundary Conditions: Land cover type and associated biophysical attributes, including FPAR, leaf area index, roughness length, albedo, etc., are all necessary to specify the state and activity of vegetation in the models. Soils, snow cover and ice data are needed by hydrological submodels.

Forcings: Near-surface meteorological conditions (temperature, humidity, wind speed), radiation fluxes and precipitation are needed by almost all land-atmosphere models. Many of the energy-water-carbon models require that the diurnal cycle be resolved in these data and that the precipitation forcing be divided into convective and large- scale fractions.

Fluxes: The energy-water-carbon and biogeochemistry models calculate the land-atmosphere exchanges of energy, water, carbon, and trace constituents and changes in equivalent storage quantities within the vegetation-soil system. For example, land surface parameterizations in GCMs typically produce time-series of evapotranspiration, soil moisture,snow and ice storage, and runoff. With the exception of some satellite-derived surface radiation budget data sets, there are no truly global data sets available that can be used to continuously validate the output from these models; the communities have to make do with temporally and spatially sparse surface-atmosphere flux data sets, which are mainly derived from field experiment data, and a few runoff records.

Within this framework, each modeling group prepared its own prioritized list of data sets. When these were analyzed and compared, it was found that there was a large overlap in the stated requirements. Table 1 lists the consolidated data needs as prioritized across the three working groups at the workshop. At the time, these high priority data sets were perceived to be unavailable or inaccessible to the modeling community. Specifically;

a. Operational meteorological agencies generate streams of 4-dimensional data assimilation (4DDA) products, including near-surface meteorology, radiation fluxes, soil moisture fields, etc., but the required information specified by the working groups was expensive and difficult to extract from the product archives.

- Very few satellite-based data products were actually available.

- Other data sets based on surface survey work (soils, topography, runoff were available but required considerable further analysis or reduction to make them directly useful to the modelers.

With regard to data accessibility, it was thought that with some effort the situation could be greatly improved. In most cases, such as the 4DDA products, soils information, topography, etc., it was thought to be more a question of institutions deciding to take on the job and committing resources to see it through, rather than the solution of difficult technical problems.

The situation with respect to data availability was different: archives of satellite data certainly existed in the form of instrument counts, exoatmospheric radiances, or in some cases atmospherically corrected surface radiances. In only a handful of cases, for example the International Satellite Cloud Climatology Project (ISCCP) cloud products and the Earth Radiation Budget Experiment (ERBE) surface (clear-sky) albedo products, were there global fields of surface or atmospheric parameters. For some of the satellite-based products specified in Table 1 (vegetation, incident PAR, insolation), the raw satellite data existed but the processing had not been carried through to the production of global data sets of physical or biophysical parameters. However, most of the necessary tools and materials for undertaking such a project were available at the time of the workshop: the data existed, many of the algorithms had been developed and tested using field experiment data, and the required data product list was defined. What was required was an initiative to bring all of these things and the appropriate scientific expertise together to actually produce the global data sets. It was repeatedly pointed out that huge resources had been expended by agencies to design and launch satellite instruments, collect and archive the observations, and conduct the necessary investigations to understand and use the data. The final step, applying recent scientific experience to produce global data sets of useful and usable parameters, was a clear priority and would be relatively cheap to execute, but had been done in only a few cases. These general assessments formed the basis for some specific recommendations.

1.5 Workshop Recommendations

Three initiatives were put forward by the workshop. These cover the immediate generation of global data sets, the improvement of methodologies and algorithms for follow-on data sets, and the improvement of communications between different elements of the Land Science community. These are discussed in turn below.

INITIATIVE I
Immediate Generation of High Priority Global Data Sets

The original 1992 workshop recommendation is restated more or less verbatim here. It is proposed that some essential global data sets could be put together within 2 years, i.e., by the summer of 1994, and released to the community. Existing or planned data management systems should be involved in this effort from the beginning. The data sets are listed in order of priority in Table 1 and are shown schematically in Figure 2. The workshop made the following recommendations for the four areas of vegetation, hydrometeorology, radiation, and soils.

Vegetation: Global, monthly data sets of vegetation-related parameters should be generated at good spatial resolution (100 x 100 km or better is preferred) and monthly time resolution. The available AVHRR data should be used as the basis of this effort and algorithms applied to calculate fields of cover type, phenology, FPAR and leaf area index.

Hydrometeorology: Near-surface meteorological data sets should be extracted from the 4-dimensional data assimilation (4DDA) streams generated by operational meteorological agencies. Specifically, near-surface temperature, humidity, wind vector, surface temperature, soil moisture content, radiation components and precipitation should all be saved. Temporal resolution should be sufficient to resolve the diurnal cycle (preferably four or more reports per day).

A number of institutions hold archives of rainfall data. A gridded product (100 x 100 km or better) is required with monthly time resolution and some information, direct or indirect, on the proportion of convective to large-scale precipitation. Runoff data is stored at the Global Runoff Data Center (GRDC) in Germany—these data should be processed to yield mm/day numbers (monthly means) for selected large catchments. This data subset would be of more direct use to modelers.

Snow and ice data are collated by NOAA and NASA in the U.S. and also by Canadian and Russian operational agencies, largely from analyses of optical satellite data and in situ observations. The temporal resolution of the data should be sufficient to resolve weather-related changes in snow extent.

Radiation: There is a strong desire to have many components of the surface radiation budget available at resolutions down to 50 x 50 km, although it is clear the community could do good work with coarser resolution (250 x 250 km) products. Again, the temporal resolution should be sufficient to resolve the diurnal cycle. ISCCP holds data sets on insolation and longwave fluxes on a 250 x 250 km grid. The continuing ERBE work provides surface albedo estimates and net surface shortwave radiation fluxes on the same scale.

Soils, Soil Moisture, and Topography: Global soils data sets with quantitative even if only best-guess, soils physics and soils chemistry information are needed. Soil texture, depth, porosity, mineralogy, and pH fields are required by some water-energy-vegetation and most biogeochemistry modelers. A data set could be quickly generated based on the Food and Agricultural Organization (FAO) global 1 ° x 1 ° data base and supporting or related information.

Soil moisture information is very useful for validating all classes of models It was recommended that the soil moisture remote sensing community be tasked with producing some global or regional products from existing sources, such as in situ observations and spaceborne microwave sensors (e.g., SSM/I), even if this turns out to give only qualitative spatial and temporal patterns of soil moisture climatology, rather than precise information at a single point under ideal retrieval conditions. (These patterns would be very useful for checking 4DDA fields and other soil moisture estimates).

Good topographic data sets are available but not easily accessible. Every effort should be made to extract the best available product from the U.S. Geological Survey (USGS) or the Defense Mapping Agency (DMA).

This recommendation framed ISLSCP Initiative I. After the 1992 workshop, an ad hoc ISLSCP Science Steering Committee supported by staff at NASA GSFC worked to put together a mutually consistent collection of data sets that would meet the needs expressed in Table 1. This effort has resulted in the issue of this collection of CDs, the contents of which are summarized in section 2. In large part, these data sets satisfy the requirements stated in Table 1, except for those specifying soil chemical properties, which hopefully will be addressed by elements of IGBP, and topography, which is being handled by a team at EROS Data Center as part of the Earth Observing System (EOS) project. It should be noted that, as requested, all the data were reformatted to a common 1 ° x 1 ° grid and cover the same period, 1987–1988.

The Initiative I CDs should be an invaluable resource for initializing, forcing, and validating all three classes of land models, see Figure 2. One example; the International Geosphere-Biosphere Project (IGBP) may use the CDs as a baseline initial condition and meteorological forcing data set for a global carbon model intercomparison exercise. Another example; the data on the CDs will be used to force offline versions of land surface parameterizations

(LSPs) to calculate more realistic global fields of hydrological variables including; evapotranspiration, soil moisture, runoff, etc. This last project is sponsored by GEWEX-ISLSCP and IGBP-BAHC. Besides these and similar applications, the data set will provide a strong starting position for global studies that will help the Land Science community prepare for the Earth Observing System data stream.

INITIATIVE II
Improved Follow-On Data Sets

The data sets specified in Initiative I were generated over a 2 year period i.e., with existing data and the available robust and simple algorithms. The resulting products go some way toward satisfying the immediate needs of the modelers and will exercise every aspect of the data-algorithm-modeler pipeline as well as (hopefully) a data system or two en route. However, it is clear that great improvements could be made over this first data release, mainly in the areas of temporal coverage, algorithm improvement, and validation. ISLSCP Initiative II has the aim of releasing an improved set of global data in 1997, which should cover the period 1986–1995.

INITIATIVE III
Improved Communications Within the Land Science Community

The workshop highlighted the extent to which related research thrusts can become separated from each other even when it is obvious that there are strong mutual scientific interests at stake. It was recognized that top-down coordination by management could provide only part of the answer. It is equally important to provide regular forums where the different communities discuss their areas of overlap on a scientist-to-scientist basis. It was observed that many recent workshops had drifted into within-discipline discussions (e.g., wish-list writing, experiment design, etc.) with little time to focus on the so-called bottleneck issues (e.g., implementation of algorithms to produce global data sets, incorporation of late-developing model needs into experiment design, etc.). Clearly, these cross-cutting issues need explicit attention.

II. Assembling the Initiative I Global Data Sets

2.1 Specification of the Properties of the Data Set

The workshop recommendations for Initiative I provided the starting point for the collection, compilation and documentation of the global data sets necessary to satisfy the requirements summarized in Table 1 and section 1.5. Immediately after the 1992 workshop, a team at NASA GSFC started to communicate with possible sources of the required data and worked to define the form of the final product.

The previous section mentioned that modelers were frequently hampered by the need to match up incongruent data sets, a task that not only wastes time but can also inject artifacts into the data so that model-to-model comparisons become less exact, depending on the type and number of regridding or interpolation operations performed on the original data. Initiative I specified the need for uniform data sets, which is interpreted to mean that the data sets should have as far as possible the same spatial and temporal resolutions, time period, and area of coverage (i.e., no spatial or temporal gaps) and be supported by uniform documentation. In principle, it should be possible to operate a land-atmosphere model continuously over the entire spatial and temporal domain of the data without encountering problems, such as missing data, in the process. Similarly, it should be possible to select any grid point and access all the necessary data to initialize and force a land- surface model over the period covered by the data set. To satisfy these requirements, the data should be spatially and temporally uniform and contiguous, and should also be uniform “vertically”; that is, a common spatial resolution allows for the “stacking” of different data sets over the same grid area, which makes for much easier 1-dimensional model operation. Taking Table 1 as a basis, the team surveyed available data sources and decided on the following attributes for the final products.

Spatial Resolution: All data were obtained or were regridded to a single 1° x 1° equal-angle grid. The sides of each grid box are specified by integer latitude/longitude lines. A single land mask was applied to all the data.

The regridding procedure was very simple: a 1° x 1° grid was laid over the source data field and area-weighted averages of the values falling within each new grid square were calculated. No smoothing or other interpolation procedures were applied to the data except where data were missing, in which case simple interpolation schemes were used. These schemes are described in detail in the data set documentation on the CDs.

Temporal Resolution: The temporal resolutions of the data sets are nested to resolve the diurnal and seasonal cycles as appropriate, see Table 2. The forcing data sets—near-surface meteorology, radiation, and precipitation—are provided as 6-hourly values so that the diurnal cycle is resolved as requested by the energy-water-carbon modelers. Most of the other data sets are monthly (e.g., vegetation attributes) or fixed (e.g., soil type). Some of the shortwave radiation data are presented as diurnally resolved, monthly means; that is, for each month, eight mean radiation fields are provided at 3-hourly GMT intervals: 0000Z, 0300Z, 0600Z and so on.

Spatial Coverage: With a few well-documented exceptions (runoff, snow cover and depth, and some of the radiation products), all the data are spatially continuous over the specified land mask. In some cases, this meant that data sets had to be interpolated spatially so as to prevent leaving holes; when this was done, a mask showing which data points were synthesized was generated. The objective was to provide a reliable data set that would allow continuous operation of land-atmosphere models without having to invoke complex procedures to deal with null data points. The exceptions are represented by data sets that are to be used for validation rather than model operation.

Temporal Coverage: The initial requirement was for 1 year's worth of data. This was extended to 2 years to provide some notion of interannual variability. It was decided to choose contiguous years for ease of model operation. The period 1987–1988 was selected, as it covers a period when many of the source data sets were simultaneously available and also covers wet (1987) and dry (1988) summer conditions in North America.

Formatting and Documentation: A variety of data formatting options were considered. It was finally decided that in order to ensure the easiest and widest possible use of the data, all of the data sets would be represented as simple ASCII files. Each global field starts at 90° N, 180° W and is read 360 grid cells toward the east before dropping down a row to start again at 89° N, 180° W; in other words, the data read like written text from the North pole and dateline southward. Data at very high latitudes are usually meaningless due to the small areas involved; in these cases, the grid cells are filled by replicating values from adjacent cells. Nulls in the data sets are represented by negative nines which are specified at the same numerical resolution as the data; e.g., the null for a 3 digit number is –999.

The documentation follows a consistent format across all the data sets, see Table 3. This has obvious advantages; after a short learning process, the user can easily target specific sections to get the desired information on any of the data sets. The documentation is also fairly detailed so that the user is not directed to external sources of information except for really indepth material on sensors or analysis techniques. References are provided.

2.2 Compilation of the Data Sets

The data sets are organized into the following categories on the CD.

Vegetation: Land Cover and Biophysics
Hydrology and Soils
Snow, Ice, and Oceans
Radiation and Clouds
Near-Surface Meteorology

The subsections below briefly review the contents of each of these data categories, further information can be found in the documentation accompanying each data set.

2.2.1 Vegetation: Land Cover and Biophysics (Table 2A)

The basis for this data set is the Normalized Difference Vegetation Index (NDVI) data set calculated from AVHRR data by Los et al. (1994) following the work of Tucker et al. (1986). These data were already in the form of a 1° x 1° monthly composited NDVI data set; i.e., no further single channel data or geometric information were available at the time. Some simple procedures were used to fill in gaps in the data set, and crude corrections were made to account for the effects of solar angle and persistent clouds to make the temporally and spatially continuous FASIR-NDVI product, see Sellers et al. (1994). The FASIR-NDVI data were used to create fields of FPAR, leaf area index, and greenness, which in turn were used to calculate monthly snow-free albedo and surface roughness fields, see Sellers et al. (1994). The land/sea mask associated with these data sets was adopted as the standard for masking the other data placed on the CD.

All of these operations, starting with the production of the FASIR-NDVI fields, require some assumptions about land cover type. DeFries and Townshend (1994) analyzed NDVI data to specify the distribution of land cover types for the world. This classification map was used to apply vegetation cover-specific algorithms for the calculation of the higher order products listed in Table 2A. The documentation for this data set also includes parameter values associated with each vegetation type as used in the SiB2 model of Sellers et al. (in prep.).

Soil background fields had to be specified as lower boundary conditions for the calculation of the snow-free albedo in Table 2A. For the most part, background (soil or litter layer) reflectances were assigned values typical of each vegetation type, as specified in DeFries and Townshend (1994), in the same way as was done by Dorman and Sellers (1989). However, this procedure resulted in some problems in sparsely-vegetated regions so ERBE data were used to estimate surface reflectances in desert areas between 45° S and 45° N, see Sellers et al. (1994).

Last, the documentation for this data set includes parameter values associated with each vegetation type as used in the SiB2 model of Sellers et al. (in prep.).

2.2.2 Hydrology and Soils(Table 2B)

The Global Precipitation Climatology Project (GPCP) reanalyzed their archive of surface rain gauge data to produce a 1° x 1° monthly precipitation product for 1987–1988. The standard land/sea mask was applied by the publication group at NASA GSFC, see Figure 3. The Global Runoff Data Center (GRDC) contributed monthly river runoff rate data for 14 basins together with information on the location of the gauges and the catchment area upstream of the gauge. The percentage of each 1° x 1° grid area covered by lakes, rivers, and marshes was obtained from data published by Cogley (1991).

The Food and Agricultural Organization (FAO) archive on soil properties has been extensively scrutinized by researchers at the University of Arizona (Sorooshian, Amer), NASA GSFC (Koster) and at NASA GISS (Zobler). These analyses were combined to create consistent global fields of soil texture, depth, and slope.

2.2.3 Snow, Ice, and Oceans (Table 2C)

NOAA NESDIS provides a weekly analysis of Northern Hemisphere snow cover from optical satellite data; the analyses are done by hand. Robinson (pers. comm.) of Rutgers University provided these data after regridding them to 1° x 1°. The U.S. Air Force assembles a monthly snow depth map based on a variety of sources including satellite data and in situ measurements carried out at reporting airfields. The NOAA National Meteorological Center (NMC) provided analyses of sea ice cover and sea surface temperature at monthly time resolution at the required 1° x 1° resolution. The land/sea mask was then applied at GSFC. Last, a fine resolution map of the land-ocean boundary was provided by the National Center for Atmospheric Research (NCAR) based on data collated by the U.S. Navy.

Global monthly fields of sea surface temperature (SST) and sea ice concentration were also included in the data set in response to requests from GCM modelers who wished to have a complete set of surface boundary conditions on the CD.

2.2.4 Radiation and Clouds (Table 2D)

Pinker and Laszlo of the University of Maryland processed the satellite data analyses held in the ISCCP archive, see Schiffer and Rossow (1985), to create five global radiation products. The 2.5° x 2.5° ISCCP data were used to generate estimates of the surface and top of the atmosphere (TOA) incident and upwelling shortwave fluxes. In addition, the surface downwelling PAR flux was also calculated. These estimates were generated every 3 hours based on GMT observing times, i.e., 0000Z, 0300Z, 0600Z, etc. To reduce the noise in their products, Pinker and Laszlo averaged the observations for each 3-hour period by month to produce a mean diurnal cycle of eight (monthly-averaged) values for each month in 1987–1988. NASA LaRC used a similar methodology to generate monthly means (not diurnally resolved) of surface incident and net shortwave and longwave radiation, and net radiation, see Darnell et al. (1992).

The ISCCP group at NASA GISS generated a series of cloud parameters from analyses of the ISCCP C2 satellite data archive. These include cloud amount, cloud top pressure, cloud optical thickness, and cloud water paths, see Rossow et al. (1991) and Rossow and Schiffer (1991).

The ERBE S4 clear-sky albedo product was generated from composites of satellite data, see Barkstrom et al. (1990). These data are provided as monthly means and do not extend beyond the solar terminator. They are likely to be dubious in persistently cloudy areas.

All the products described above were originally generated on a 2.5° x 2.5° equal-area grid, which was reprocessed at NASA GSFC onto the 1° x 1° grid used by ISLSCP, see Figure 3. However, the standard land/sea mask was not applied to any of these data sets. The diurnally resolved data of Pinker and Laszlo had some“holes” in it due to gaps in some of the geostationary satellite data records used to create the ISCCP product. These “holes” were patched using a simple temporal interpolation technique that made use of solar angle information; the patched areas are flagged in the final 1° x 1° product. Some of the NASA GISS cloud product fields are also discontinuous; in particular, there are gaps close to and a complete lack of data above the solar terminator for some of the fields.

The radiation and clouds data are intended to be used as follows. The University of Maryland and NASA LaRC products are useful for forcing models; these are probably the best current estimates that we have for global surface radiation fluxes. We have further used these data in combination with GCM output to synthesize estimates of the downwelling shortwave and longwave fluxes every 6 hours, see next section. The ISCCP cloud products may be useful for testing atmospheric radiation models. The ERBE clear-sky albedo product may be useful for validating model-generated fields, in particular the effects of snow when combined with the snow-free albedo fields described in section 2.2.1. (Table 2A).

2.2.5 Near-Surface Meteorology (Table 2E)

The bulk of the near-surface meteorological products on the CD were extracted from the ECMWF operational forecast analysis archive. The data set consists of time-series of meteorological variables at 6-hourly intervals (0000Z, 0600Z, 1200Z, 1800Z) and monthly 6-hourly averages of these and many other diagnostic and prognostic variables. The meteorological variables that are required to force land-atmosphere models with resolved diurnal cycles were extracted from this stream: surface pressure, air temperature, dew point, and wind speed (magnitude). It was also desired to have incident shortwave and longwave radiation fluxes and precipitation rates, preferably broken into large-scale and convective components, at the same 6-hourly temporal resolution. However, the ECMWF output did not contain precipitation rates, and their radiation flux estimates were thought to be biased due to a systematic underestimation of cloud cover by the model version used to generate these products. To fill the gap, the NASA GSFC team generated hybrid radiation products: the time-series of ECMWF estimates of surface shortwave and longwave fluxes were used to divide up the NASA LaRC satellite-based monthly radiation fluxes into 6-hourly intervals. This resulted in the synthesis of 6-hourly incident shortwave and longwave fluxes that add up to match the NASA LaRC monthly means. A similar procedure was used at NMC by Mitchell to generate 6-hourly estimates of precipitation. Six-hourly total and convective precipitation fields from the 4DDA-based NMC Reanalysis Project (Kalnay and Jenne, 1991) were used to partition the observed monthly GPCP precipitation products into 6-hourly time series of estimated total and convective precipitation, wherein the total precipitation added up to match the GPCP monthly totals. In this procedure, a screening was applied, based on the FGGE daily rainfall data, following Liston et al. (1993), to better reproduce the observed frequency of measurable daily rainfall. All of this effort has resulted in a temporally and spatially consistent meteorological forcings data set with a 6-hourly timestep, see Table 2E(iii).

These quantities, and some others that were held on the ECMWF record at 6-hourly resolution, were processed to provide monthly 6-hourly mean products and statistics, see Table 2E(ii). Of particular interest are the ECMWF-generated estimates of the surface radiation and heat fluxes.

Monthly mean fields and associated statistics of some of the prescribed or initial fields of surface boundary conditions used by ECMWF are listed in Table 2E(i). These fields were generated from a variety of sources (see the documentation) and are not recommended as initialization or boundary condition fields for current modelers; they are provided as information to help users understand what assumptions were made in generating the forcing fields in Table 2E(iii). For example, the ECMWF snow-free albedo field is based on the products of Dorman and Sellers (1989), which are thought to be less accurate than the satellite-data based products described in Section 2.2.1.

All the ECMWF fields were converted from grid cell corner point values to values representative of the entire (ECMWF) grid cell; these were then converted to the ISLSCP 1° x 1° grid and the land/sea mask was applied, see Figure 3.

In summary, a combination of products from operational meteorological agencies (ECMWF and NMC); satellite-data based radiation estimates (NASA LaRC); and global surface rain gauge analyses (GPCP) have been used to generate time-series of the required near-surface model forcings for the period 1987–1988 at a 6-hourly time resolution. These are supported by a range of ancillary time-averaged quantities that may be useful for forcing models that run on a monthly timestep.

2.3 Peer Review

A peer review process was organized by Kerr and Meeson to ensure the quality of the data and documentation to be placed on the CDs. In the first stage, the documentation was reviewed by individuals familiar with the data sets but not directly involved in writing the documentation. The intent of this review was twofold; first, to provide a "second opinion" and second, to ensure the accuracy and clarity of the documentation. The reviewers were asked to identify subtle as well as major inaccuracies or gaps that only someone familiar with the data set would know. To provide a uniform and consistent review of the documents, a set of document review guidelines and a response form were drafted and sent with the documents to all document reviewers. Comments or corrections received from these reviewers were addressed and incorporated into documentation before it moved on to the second stage of the review procedure.

In the second stage, reviewers were sent both the revised documentation and the data, and were asked to examine them using a common set of criteria. These criteria focused on the identification of errors or inaccuracies within the data and related documentation. The reviewers were selected for their general familiarity with the type of data that they were to review. This stage of the review process was completed in two workshops that focused on the mutual consistency of the data sets and documentation. The findings of these workshops and the individual data reviews are summarized in the paper of Kerr et al., also reproduced on this CD.

III. Summary

The Initiative I data sets should provide modelers with many of the fields required to prescribe boundary conditions, and to initialize and force a wide range of land-biosphere-atmosphere models. All of the data have been processed to the same spatial resolution (1° x 1°), using the same land/sea mask and steps have been taken to ensure spatial and temporal continuity of the data. The data sets cover the period 1987–1988 at 1-monthly time resolution for most of the seasonally varying quantities and at 6-hourly resolution for the near-surface meteorological and radiative forcings.

ISLSCP Initiative II aims to improve on this effort by covering a longer time period (1986–1995), at higher spatial resolution (0.5° x 0.5°), using superior data sources and algorithms where possible. In addition, GEWEX-ISLSCP and other organizations, for example IGBP-BAHC, are pursuing approaches for collating validation data sets to check the Initiative II data sets at a few times and places embedded within these global data sets.

IV. References

Barkstrom, B.R., E.F. Harrison, and R.B. Lee (1990). Earth Radiation Budget Experiment, preliminary seasonal results. *EOS Transactions.* American Geophysical Union. 71, February 27.

Betts, A.K., J.H. Ball, A.C.M. Beljaars, M.J. Miller, and P.Viterbo (1994). Coupling between land-surface boundary-layer parameterizations and rainfall on local and regional scales: Lessons from the wet summer of 1993. *Fifth Conference on Global Change Studies: Amer. Meteor. Society Proceedings.* 74th Annual Meeting, Nashville, TN, Jan. 23–28, 1994.

Bonan, G.B. (1995). Land-atmosphere interactions for climate system models: coupling biophysical, biogeochemical, and ecosystem dynamical processes. *Rem. Sens. Env.* 51:1:57–73.

Cogley, J.G. (1991). GGHYDRO-global hydrographic data, release 2. Available from the author at Trent University, Ontario, Canada.

Darnell, W.L., W.F. Staylor, S.K. Gupta, N.A. Ritchey, and A.C. Wilber (1992). Seasonal variation of surface radiation budget derived from ISCCP-C1 data. *J. Geophys. Res.* 97:15741–15760.

DeFries, R.S., and J.R.G. Townshend (1994). NDVI-derived land cover classification at global scales. *I. J. of Remote Sensing.* 15:17:3567–3586.

Dickinson, R.E. (1995). Land processes in climate models. *Rem. Sens. Env.* 51:1:27–38.

Dorman, J.L., and P.J. Sellers (1989). A global climatology of albedo, roughness length, and stomatal resistance for atmospheric general circulation models as represented by the Simple Biosphere Model (SiB). *J. Appl. Met.* 28:9:833–855.

Field, C.B., C.M. Malmstrom, J.T. Randerson (1995). Ecosystem net primary production: Combining ecology and remote sensing. *Rem. Sens. Env.* 51:1:74–88.

Kalnay, E., and R. Jenne (1991). Summary of the NMC/NCAR reanalysis. *Bull. Amer. Meteor. Soc.* 72:897–1904.

Liston, G.E., Y.C. Sud, and G. Walker (1993). Design of a global soil moisture initialization procedure for the Simple Biosphere model. *NASA Tech. Memo. 104590.* Goddard Space Flight Center, Greenbelt, MD.

Los, S.O., C.O. Justice, and C.J. Tucker (1994). A 1° x 1° global NDVI data set for climate studies derived from the GIMMS continental NDVI data. *I. J. of Remote Sensing.* 15:3493–3518.

Noilhan, J., P. Bougeault, B. Bretl, and P. LaCarrere (1991). An example of spatial integration of a land surface parameterization in meso-beta scale model. In *Land Surface Evaporation.* Eds Schmugge and Andre. Springer-Verlag, New York. 383–402.

Rossow, W.B., L.C. Garder, P.J. Lu, and A. Walker (1991). International Satellite Cloud Climatology Project (ISLSCP): documentation of cloud data. Tech. Doc. WMO/TD No. 266 (revised). World Meteorological Organization. Geneva. 76 p. plus three appendices.

Rossow, W.B., and R.A. Schiffer (1991). ISCCP cloud data products. *Bull. Amer. Meteor. Soc.* 72:2–20.

Sato, N., P.J. Sellers, D.A. Randall, E.K. Schneider, J. Kinter III, J. Shukla, Y-T Hou, and E. Albertazzi (1989). Effects of implementing the simple biosphere model (SiB) in a general circulation model. *J. Atmos. Sci.* 46:18:2757–2782.

Schiffer, R.A., and W.B. Rossow (1985). ISCCP Global Radiance Data Set. A new resource for climate research. *Bull. Am. Meteorol. Soc.* 66:1498–1505.

Schimel, D.S. (1995). Terrestrial biogeochemical cycles: global estimates with remote sensing. *Rem. Sens. Env.* 51:1:49–56.

Sellers, P.J., D.A. Randall, C.J. Collatz, J.A. Berry, C.B. Field, D.A. Dazlich, C. Zhang, and G.D. Collelo (in prep.). A revised land surface parameterization (SiB2) for atmospheric GCMs. Part 1: Model formulation. Submitted to *J. of Climate.*

Sellers, P.J., B.W. Meeson, F.G. Hall, G. Asrar, R.E. Murphy, R.A. Schiffer, F.P. Bretherton, R.E. Dickinson, R.G. Ellingson, C.B. Field, K.F. Huemmrich, C.O. Justice, J.M. Melack, N.T. Roulet, D.S. Schimel, and P.D. Try (1995). Remote sensing of the land surface for studies of global change: models- algorithms-experiments. *Rem. Sens. Env.* 15:1:3–26.

Sellers, P.J., S.O. Los, C.J. Tucker, C.O. Justice, D.A. Dazlich, G.J. Collatz, and D.A. Randall (1994). A global 1° x 1° NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *I. J. Remote Sensing.* 15:7:3519–3545.

Sellers, P.J., S.I. Rasool, and H-J. Bolle (1990). A review of satellite data algorithms for studies of the land surface. *Bull. Amer. Met. Soc.* 71:10:1429–1447.

Tans, P.P., I.Y. Fung, and T. Takahashi (1990). Observational constraints on the global atmospheric carbon dioxide budget. *Science.* 247:1431–1438.

Tucker, C.J., I.Y. Fung, C.D. Keeling and R.H. Gammon (1986). Relationship between atmospheric carbon dioxide variations and a satellite-derived vegetation index. *Nature.* 319:195–199.

Acknowledgements

Many people and organizations worked hard to turn Initiative I into a reality.

First, special thanks are due to Ghassem Asrar of NASA Headquarters; David Schimel of CSMP; TERRA Laboratory, a consortium of the USDA's Agriculture Resource Service (Steve Rawling and Don DeCoursey); the USDA's Forest Service (Doug Fox); and the USGS (Ray Watts) for providing financial support for the 1992 ISLSCP Workshop that started the activity. The bulk of the funding came from the EOS program at NASA HQ. Dr. Asrar is particularly thanked for his unflagging moral support of this effort from start to finish.

Next, financial support for the data compilation and production phase of the CDs was provided by Drs. Bob Murphy and Tony Janetos of NASA HQ. Encouragement was provided by the chair of WCRP-GEWEX, Dr. Moustafa Chahine and the Director of WCRP, Dr. Pierre Morel. All are warmly thanked. The support of NASA Headquarters, Office of Mission to Planet Earth, the Operations, Data and Information Systems Division, and the Science Division are gratefully acknowledged. The authors would also like to thank the EOS-DIS, Goddard Distributed Active Archive Center (GSFC DAAC) for their support of this work.

In addition to review work done by the authors, the data sets and documentation were reviewed by volunteers in the community: Nigel Arnell, Lahouari Bounoua, Peter Briggs, Gerard Dedieu, Bob Dickinson, Han Dolman, John Gash, Barry Goodison, Fred Huemmrich, Alfredo Huete, John Janoviak, Jenny Lean, Jean-Claude Menaut, Joel Noilhan, Michael Raupach, Chet Ropelewski, Bill Rossow, Steve Running, T.R.E. Thompson, Anne Walker, Ivan Wright, YongKang Xue.

In many cases, the donors of the data sets gave up a great deal of their time in addition to the data sets; all are thanked for the help. The list of individuals includes:

Vegetation: Land Cover and Biophysics

Jim Tucker, Chris Justice, Sietse Los, Piers Sellers, Don Dazlich, Jim Collatz, Nazmi El Saleous, Ruth DeFries, John Townshend, Ed Harrison.

Hydrology and Soils

Wolfgang Grabs, Arnold Gruber, J.G. Cogley, Saud Amer, Soroosh Sorooshian, Leonard Zobler, Norman Bliss, Dan Braithwaite, Randy Koster, Bruno Rudolf, Udo Schneider, Paul Try.

Snow, Ice and Oceans

Dudley Foster, David Robinson, Jay Wright, Richard Reynolds, Bob Grumbine.

Radiation and Clouds
Rachel Pinker, Istvan Laszlo, Wayne Darnell, Charles Whitlock, Bruce Barkstrom, Ed Harrison, Bill Rossow, Bob Schiffer, Chris Brest, W. F. Staylor.

Near-Surface Meteorology
Tony Hollingsworth, Horst Bottger, Ken Mitchell, Ying Lin.

The organizations who own or sponsored the collation of the data are listed in the data documentation and on the cover of the CDs. Special mention goes to

NASA GSFC Branches 923 and 974
NASA GSFC DAAC 902.2
WCRP-GEWEX elements: GRDC, GPCP, ISCCP; and IGBP-BAHC
Universities: Maryland, Arizona, Trent, Rutgers
Scientific Research Centers: NOAA NMC, ECMWF, NASA GISS, NASA LaRC, USAF ETAC, EROS Data Center, USGS.

The ISLSCP Steering Committee are warmly thanked for putting up with many reviews of the progress of the CDs and for providing valuable guidance to the Initiative I team throughout the long process of assembling the data sets. The steering committee and invited experts included: Nigel Arnell, Dennis Baldocchi, Alan Betts, Josef Cihlar, Ray Desjardins, Robert Dickinson, Chris Field, Barry Goodison, Forrest Hall, Chris Justice, Pavel Kabat, Yann Kerr, Nobuo Sato, Jerry Melillo, Carlos Nobre, John Norman, Michael Raupach, Steve Running, Piers Sellers, Jim Shuttleworth, Soroosh Sorooshian, Jim Wallace; (ex-officio) Ichtiaque Rasool, John Townshend, Moustafa Chahine, Jim Dodge, Paul Try, Ghassem Asrar, Tony Janetos, Bob Murphy, Bob Schiffer, Mike Coughlan, Pierre Morel.

Thanks go to Dawn Erlich for arranging ISLSCP meeting logistics. Last but by no means least, thanks go to Laura Blasingame and Valerie McElroy for typing and editing this paper.

Acronyms

4DDA	4-Dimensional Data Assimilation
ASCII	American Standard Code for Information Interchange
AVHRR	Advanced Very High Resolution Radiometer
BAHC	Biospheric Aspects of the Hydrological Cycle (IGBP Core Project)
BGC	Biogeochemistry
CD	Compact Disc
CD-ROM	Compact Disc-Read Only Memory
CSMP	Climate Simulation Modeling Project
DAAC	Distributed Active Archive Center
DMA	Defense Mapping Agency
ECMWF	European Center for Medium-Range Weather Forecasts
EOS	Earth Observing System
ERBE	Earth Radiation Budget Experiment
ERS-1	European Research Satellite-1
ESA	European Space Agency
FASIR	Fourier-Adjusted, Solar Zenith Angle Corrected, Interpolated and Reconstructed Data
FAO	Food and Agriculture Organization (UN)
FGGE	First GARP Global Experiment
FIFE	First ISLSCP Field Experiment
FPAR	Fraction of PAR Absorbed by the Vegetation Canopy
GCM	General Circulation Model (of the Atmosphere)
GEWEX	Global Energy and Water Cycle Experiment
GISS	Goddard Institute for Space Studies (NASA)
GOES	Geostationary Operational Environmental Satellite
GMT	Greenwich Mean Time
GPCP	Global Precipitation Climatology Project
GRDC	Global Runoff Data Center
GSFC	Goddard Space Flight Center (NASA)
HAPEX	Hydrology-Atmosphere Pilot Experiment
IGBP	International Geosphere-Biosphere Project
IGBP-DIS	IGBP-Data and Information System
ISCCP	International Satellite Cloud Climatology Project
ISLSCP	International Satellite Land Surface Climatology Project
JMA	Japanese Meteorological Agency

LAI	Leaf Area Index
LaRC	Langley Research Center (NASA)
LSP	Land Surface Parameterization
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NDVI	Normalized Difference Vegetation Index
NESDIS	NOAA Environmental Satellite Data and Information Service
NMC	National Meteorological Center
NOAA	National Oceanic and Atmospheric Administration
PAR	Photosynthetically Active Radiation
SAR	Synthetic Aperture Radar
SPOT	Systeme Probatoire pour L'Observation de la Terre
SRB	Surface Radiation Budget
SSM/I	Special Sensor Microwave Imager
SST	Sea Surface Temperature
TOMS	Total Ozone Mapping Spectrometer
UN	United Nations
USGS	United States Geological Survey
WCRP	World Climate Research Program

Tables

Table 1: Recommendations From the 1992 ISLSCP Workshop

Consolidated, Prioritized Data Needs Across the Science Areas:
Water-Energy-Carbon, Biogeochemistry, Ecological Structure, and Function

Data Fields	Domain	Resolution		Source Methodology	Action
		Spatial	Temporal		
Vegetation (cover type, phenology, disturbance, LAI, FPAR, etc.)	Regional and Global	50 x 50 km to 1x1 km	Monthly	(1) AVHRR (2) Landsat, Spot	(1) Use an existing AVHRR product for now. (2) Support 1 x 1 km land surface data set effort. (3) Revitalize efforts to correct data and apply algorithms to define biophysical parameters.
Near-Surface Meteorology	Global	50 x 50 km	Diurnal cycle, Monthly means	NMC, ECMWF, JMA; 4DDA, & observations	(1) Initiate work to process 4DDA products into usable data sets.
Precipitation	Global	100 x 100 km	Monthly means & selected days	WCRP–GPCP, Operational Met. Agencies; Surface data, Thermal IR, 4DDA	(1) Implement NMC workshop to analyze surface network data. (2) Check that the above is linked to WCRP-GPCP. (3) Provide resources for gridding data if necessary.
Radiation Fluxes (SW & LW, incoming & outgoing, PAR incoming)	Global	250 x 250 km to 50 x 50 km	Diurnal cycle, Monthly means	GOES, METEOSAT, ERBE, AVHRR, TOMS; ISCCP, ESA, NASA analyses	(1) Define interested communities, dialogue with ISCCP. (2) Check regressions using Pathfinder data. (3) Validate against long-term data.
Soil Physics: texture, depth, porosity. Chemistry: pH mineralogy	Global	100 x 100 km to 1 x 1 km	Once	FAO product & supporting material; New initiatives, notably IGBP	(1) Assign soil physics parameters to the FAO soil descriptor fields for now. (2) Support new initiative, and encourage early deliveries.
Topography	Global	10 x 10 km to 1 km or better	Once	USGS, DMA, ERS-1	(1) Support efforts to release all data from DMA. (2) Check across data sets for consistency.

—more—

Table 1: Recommendations From the 1992 ISLSCP Workshop, continued

Consolidated, Prioritized Data Needs Across the Science Areas:
Water-Energy-Carbon, Biogeochemistry, Ecological Structure, and Function

Data Fields	Domain	Resolution		Source Methodology	Action
		Spatial	Temporal		
Runoff	Regional to Global	Catchmnt grid formats 50 x 50 km	Monthly	Global Runoff Data Center (GRDC) in Germany	(1) Strong encouragement to GRDC in Germany, enlist WMO support. (2) Encourage continuous updating of the data set; gridding & averaged products.
Snow and Ice	Regional to Global	25 x 25 km	Monthly	NOAA, NASA, Russian, and Canadian agencies; SSM/I and surface observations	(1) Apply existing techniques. (2) Develop and apply improved algorithms and international communications links. (3) Investigate use of SAR.

Table 2: Data Sets on the CD

Temporal Resolutions are Given in the Right-Hand Column

Notes:

- "Monthly 3-hourly" refers to values that are monthly means of 3-hourly data. Thus, all the 0000Z values for a month are averaged into a single value, also the 0300Z values, etc.
- The snow-free albedo data set in section A is based on NDVI fields and a model calculation, the albedo field in section D is based on ERBE data, and the fields in section E originate from a survey of in-situ work.
- The documentation for the vegetation class data in section A includes vegetation morphological and physiological parameters associated with each vegetation type in the SiB2 model of Sellers et al. (in prep.).

A. VEGETATION: LAND COVER AND BIOPHYSICS—(NASA/GSFC, CSU, U. Maryland)

NDVI, FASIR-NDVI	Monthly	Background (soil/litter) reflectance (Vis, NIR)	Fixed
FPAR, LAI, Greenness	Monthly	Vegetation class	Fixed
Surface roughness, snow-free albedo	Monthly		

B. HYDROLOGY AND SOILS
(GPCP, GRDC, U. Arizona, Trent U., NCAR, FAO, NASA GSFC, NASA GISS)

Precipitation (GPCP)	Monthly	Lake, river, marsh cover percentage	Fixed
River runoff (GRDC; 14 basins)	Monthly	Soil texture, depth, slope	Fixed

—more—

Table 2: Data Sets on the CD, continued

Temporal Resolutions are Given in the Right-Hand Column

C. SNOW, ICE AND OCEANS (NOAA/NESDIS, Rutgers U., USAF, NOAA/NMC, US Navy, NCAR)	
Snow cover; depth ....Monthly Sea ice, SST ....Monthly	Land-ocean boundary ....Fixed
D. RADIATION AND CLOUDS (U. of Maryland, NASA/LaRC, ISCCP, NASA/GISS)	
Surface and TOA incoming and outgoing shortwave ....Monthly 3-hourly Surface incoming PAR fluxes ....Monthly Surface incoming shortwave and longwave radiation fluxes ....Monthly	Surface net shortwave, net longwave, net radiation fluxes ....Monthly Cloud amount, cloud top pressure ....Monthly Optical thickness, water path .... Monthly Clear-sky albedo (ERBE) ....Monthly
E. NEAR-SURFACE METEOROLOGY (ECMWF, NASA/GSFC, NOAA/NMC, NASA/LaRC, GPCP)	
(i) Prescribed/diagnostic fields Soil moisture .... Monthly Deep soil temperature and soil wetness .... Monthly Snow depth ....Monthly Albedo, surface roughness .... Fixed (ii) Monthly 6-hourly forcing fields Surface pressure, air temperature, dew point ....Monthly 6-hourly Surface temperature .... Monthly 6-hourly Mean sea level pressure ....Monthly 6-hourly u, v wind speed and stress ....Monthly 6-hourly	Surface sensible and latent heat fluxes ....Monthly 6-hourly Net surface and TOA shortwave, longwave fluxes ....Monthly 6-hourly (iii) Diurnally-resolved (6-hourly) forcing fields Surface pressure, air temperature, dew point, wind speed ....6-hourly Hybrid longwave and shortwave incoming radiation fluxes ....6-hourly Hybrid total precipitation and convective precipitation 6-hourly

Table 3: Standardized Documentation Format for the Initiative I Data Sets

1. TITLE
   - 1.1 Data Set Identification
   - 1.2 Data Base Table Name
   - 1.3 CD-ROM File Name
   - 1.4 Revision Date of This Document
2. INVESTIGATOR(S)
   - 2.1 Investigator(s) Name and Title
   - 2.2 Title of Investigation
   - 2.3 Contacts (for Data Production Information)
   - 2.4 Requested Form of Acknowledgment
3. INTRODUCTION
   - 3.1 Objective/Purpose
   - 3.2 Summary of Parameters
   - 3.3 Discussion
4. THEORY OF MEASUREMENTS
5. EQUIPMENT
   - 5.1 Instrument Description
   - 5.2 Calibration

—more—

Table 3: Standardized Documentation Format for the Initiative I Data Sets, continued

6. PROCEDURE
   - 6.1 Data Acquisition Methods
   - 6.2 Spatial Characteristics
   - 6.3 Temporal Characteristics

7. OBSERVATIONS
   - 7.1 Field Notes

8. DATA DESCRIPTION
   - 8.1 Table Definition With Comments
   - 8.2 Type of Data (Parameters, Units, Range)
   - 8.3 Sample Data Record
   - 8.4 Data Format
   - 8.5 Related Data Sets

9. DATA MANIPULATION
   - 9.1 Formulas
   - 9.2 Data Processing Sequence
   - 9.3 Calculations
   - 9.4 Graphs and Plots

10. ERRORS
    - 10.1 Sources of Error
    - 10.2 Quality Assessment

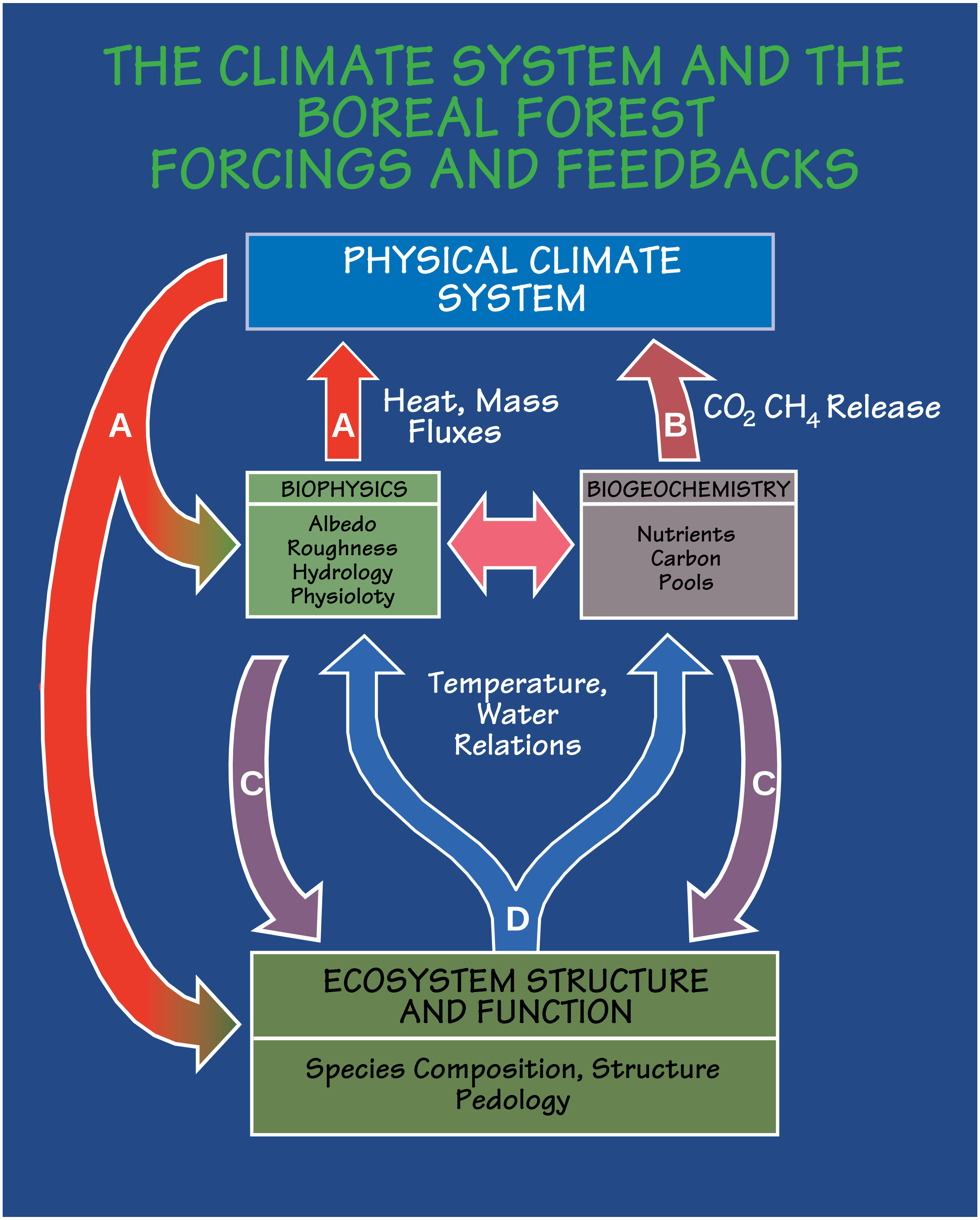
11. NOTES
    - 11.1 Known Problems With the Data
    - 11.2 Usage Guidance
    - 11.3 Other Relevant Information

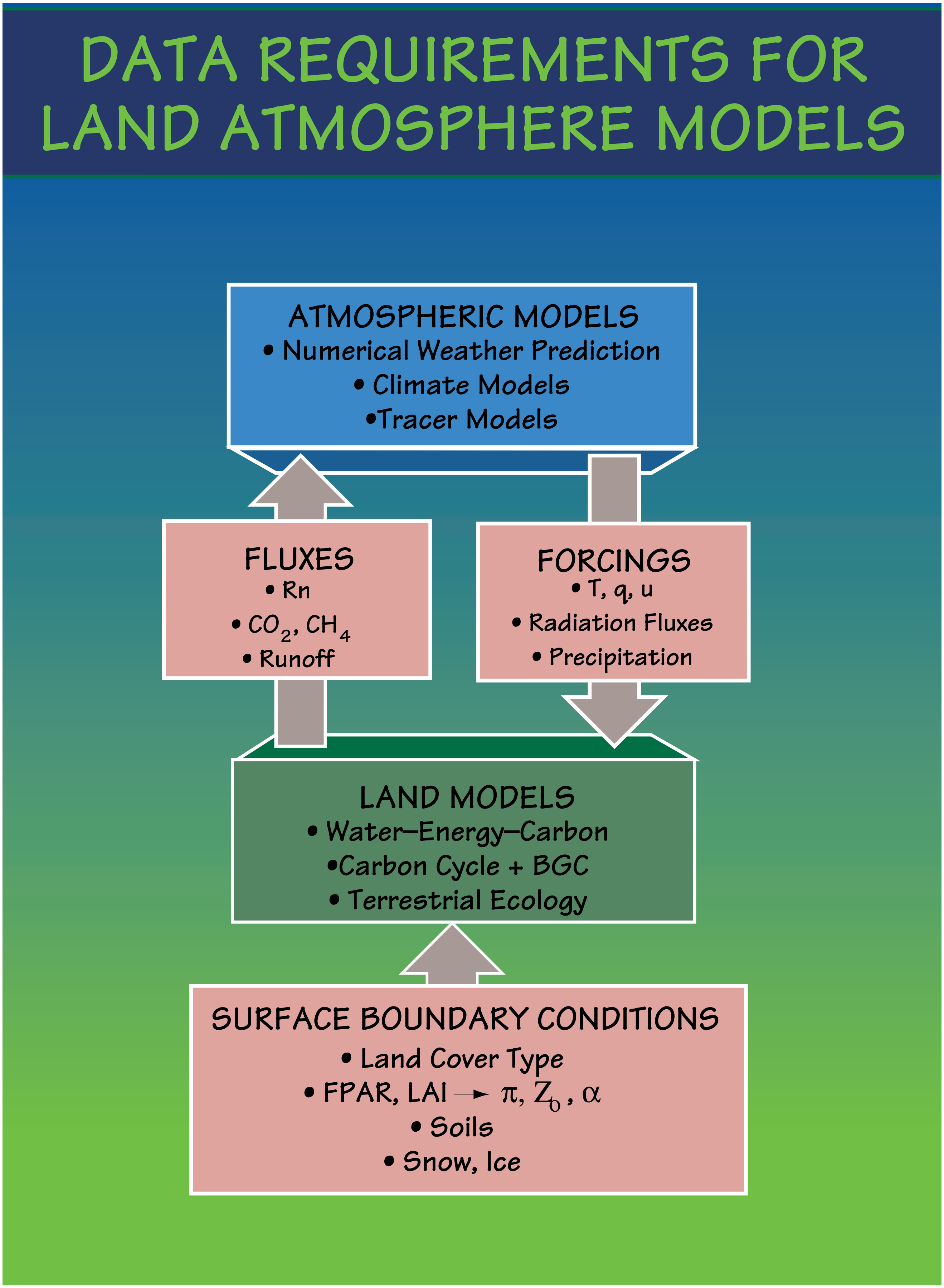
12. REFERENCES
    - 12.1 Satellite, Instrument, Data Processing Documentation
    - 12.2 Journal Articles and Study Reports
    - 12.3 Archive, DBMS Usage Documentation

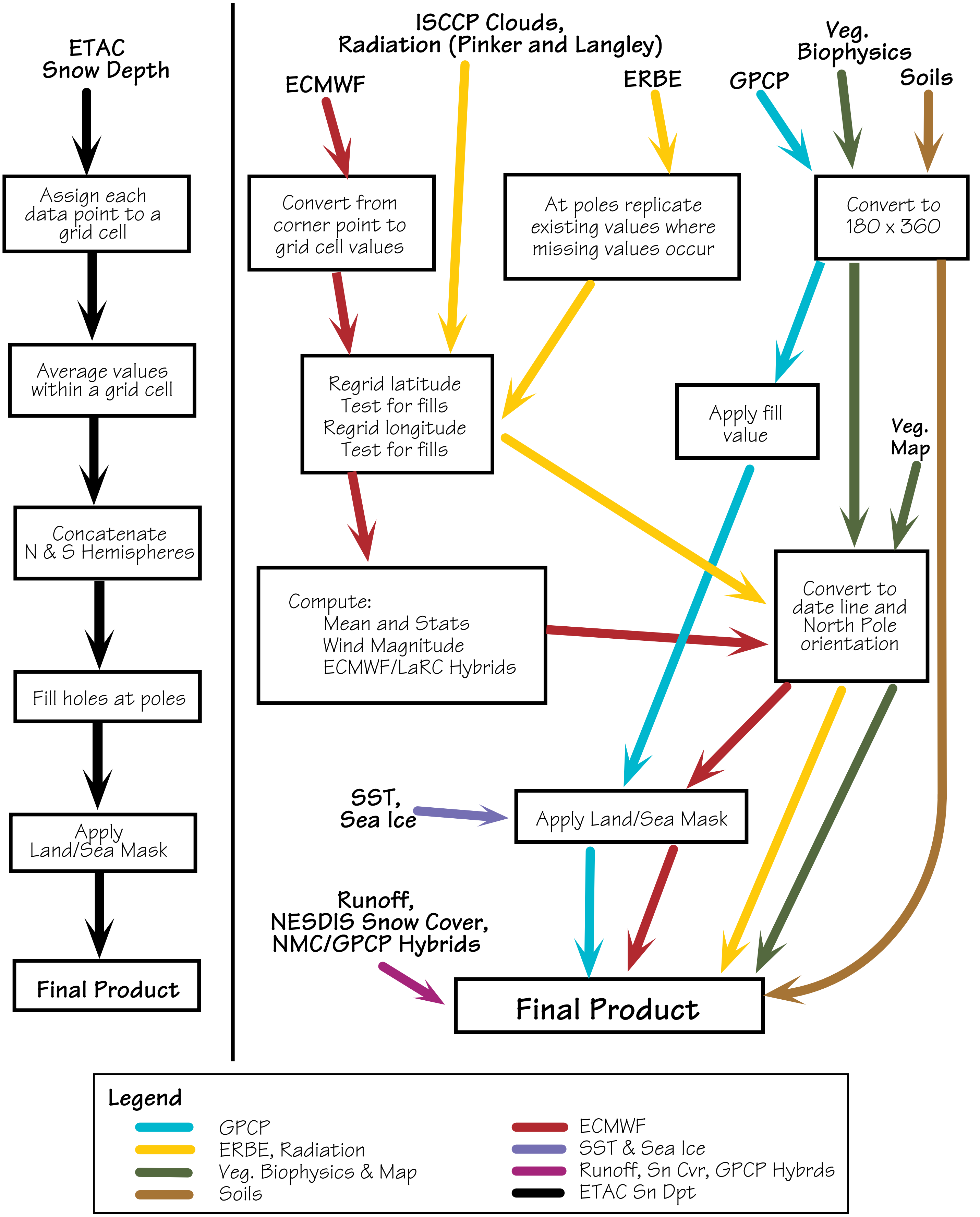
13. DATA ACCESS
    - 13.1 Contacts for Archive, Data Access Information
    - 13.2 Archive Identification
    - 13.3 Procedures for Obtaining Data
    - 13.4 Archive, Status, Plans

14. OUTPUT PRODUCTS AND AVAILABILITY
    - 14.1 Tape Products
    - 14.2 Film Products
    - 14.3 Other Products

Figures


THE CLIMATE SYSTEM AND THE BOREAL FOREST
FORCINGS AND FEEDBACKS
PHYSICAL CLIMATE SYSTEM
A
A
Heat, Mass Fluxes
B
CO_2 CH_4 Release
BIOPHYSICS
Albedo
Roughness
Hydrology
Physioloty
BIOGEOCHEMISTRY
Nutrients
Carbon
Pools
Temperature, Water Relations
C
C
D
ECOSYSTEM STRUCTURE AND FUNCTION
Species Composition, Structure
Pedology


DATA REQUIREMENTS FOR LAND ATMOSPHERE MODELS
ATMOSPHERIC MODELS
• Numerical Weather Prediction
• Climate Models
•Tracer Models
FLUXES
• Rn
• CO_2, CH_4
• Runoff
FORCINGS
• T, q, u
• Radiation Fluxes
• Precipitation
LAND MODELS
• Water–Energy–Carbon
•Carbon Cycle + BGC
• Terrestrial Ecology
SURFACE BOUNDARY CONDITIONS
• Land Cover Type
• FPAR, LAI $\rightarrow$ π, Z_0, α
• Soils
• Snow, Ice


ETAC
Snow Depth
Assign each data point to a grid cell
Average values within a grid cell
Concatenate N & S Hemispheres
Fill holes at poles
Apply Land/Sea Mask
Final Product
ISCCP Clouds, Radiation (Pinker and Langley)
ECMWF
ERBE
GPCP
Veg. Biophysics
Soils
Convert from corner point to grid cell values
At poles replicate existing values where missing values occur
Convert to 180 x 360
Regrid latitude
Test for fills
Regrid longitude
Test for fills
Apply fill value
Veg. Map
Compute:
Mean and Stats
Wind Magnitude
ECMWF/LaRC Hybrids
Convert to date line and North Pole orientation
SST, Sea Ice
Apply Land/Sea Mask
Runoff, NESDIS Snow Cover, NMC/GPCP Hybrids
Final Product
Legend
GPCP
ERBE, Radiation
Veg. Biophysics & Map
Soils
ECMWF
SST & Sea Ice
Runoff, Sn Cvr, GPCP Hybrds
ETAC Sn Dpt